

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problems Mailbox.**



PATENT ABSTRACTS OF JAPAN

(11) Publication number: **10187752 A**(43) Date of publication of application: **21 . 07 . 98**

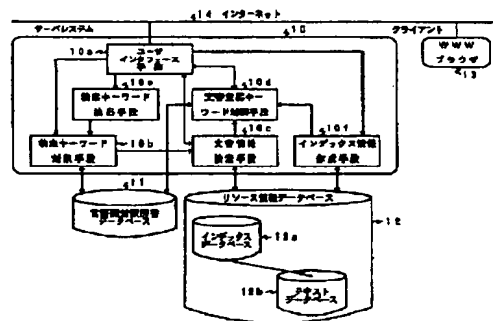
(51) Int. Cl. **G06F 17/30**
G06F 17/28

(21) Application number: **08355492**(22) Date of filing: **24 . 12 . 96**(71) Applicant: **KOKUSAI DENSHIN DENWA CO LTD <KDD>**(72) Inventor: **SUZUKI MASAMI
HASHIMOTO KAZUO****(54) INTER-LANGUAGE INFORMATION RETRIEVAL
BACKUP SYSTEM****(57) Abstract:**

PROBLEM TO BE SOLVED: To provide a retriever with a backup information for easy and accurate retrieval when document information provided in a language other than a desired language is retrieved.

SOLUTION: The inter-language information retrieval backup system is provided in a network, to which computers are connected, and equipped with a server system 10 having an inter-language equivalent dictionary data base 11 which has a dictionary for translating a key word among different kind of languages, a document main key word translating means 10d, which translates a main key word of the source language of requested document information into a desired language by using the inter-language equivalent dictionary data base 11, and a user interface means 10a which enhances and answers the document information so that the main key word translated at the same time with the document information is displayed as a backup information frame.

COPYRIGHT: (C)1998,JPO



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-187752

(43) 公開日 平成10年(1998) 7月21日

(51) Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 17/30 17/28			G 0 6 F 15/403 15/38 15/40	3 2 0 C Q 3 1 0 C 3 1 0 F 3 7 0 A
審査請求 未請求 請求項の数12 F D (全 11 頁) 最終頁に続く				

(21) 出願番号 特願平8-355492

(22) 出願日 平成8年(1996)12月24日

特許法第30条第1項適用申請有り 1996年6月28日～6月29日 開催の「PROCEEDINGS OF THE WORKSHOP ON INFORMATION RETRIEVAL WITH ORIENTAL LANGUAGES」において文書をもって発表

(71) 出願人 000001214

国際電信電話株式会社

東京都新宿区西新宿2丁目3番2号

(72) 発明者 鈴木 雅実

東京都新宿区西新宿2丁目3番2号国際電信電話株式会社内

(72) 発明者 橋本 和夫

東京都新宿区西新宿2丁目3番2号国際電信電話株式会社内

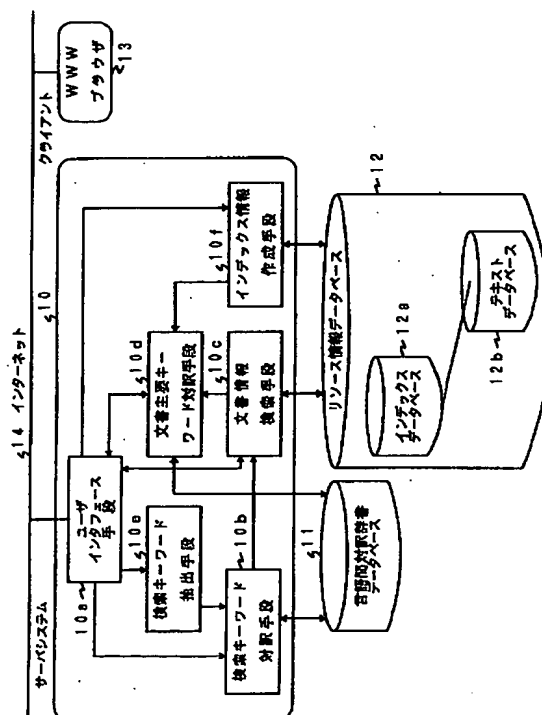
(74) 代理人 弁理士 山本 恵一

(54) 【発明の名称】 言語間情報検索支援システム

(57) 【要約】

【課題】 所望の言語以外の言語で提供された文書情報を検索する際に、容易かつ的確な検索を行うための支援情報を検索者に提供する。

【解決手段】 複数のコンピュータが接続されたネットワークに設けられており、複数の種類の言語間でキーワードを対訳するための辞書を有する言語間対訳辞書データベースと、要求された文書情報の原言語の主要キーワードを前記言語間対訳辞書データベースを用いて所望の言語に対訳する文書主要キーワード対訳手段、及び該文書情報と同時に対訳された主要キーワードを支援情報フレームとして表示するように該文書情報をエンハンスして応答するユーザインタフェース手段を有するサービスシステムとを備えている言語間情報検索支援システムである。



【特許請求の範囲】

【請求項1】 複数のコンピュータが接続されたネットワークに設けられており、
複数の種類の言語間でキーワードを対訳するための辞書を有する言語間対訳辞書データベースと、
要求された文書情報の原言語の主要キーワードを前記言語間対訳辞書データベースを用いて所望の言語に対訳する文書主要キーワード対訳手段、及び該文書情報と同時に
対訳された主要キーワードを支援情報フレームとして
表示するように該文書情報をエンハンスして応答するユーザインタフェース手段を有するサーバシステムとを備えていることを特徴とする言語間情報検索支援システム。

【請求項2】 前記サーバシステムは、更に、前記文書情報のテキストリソース全体から原言語の前記主要キーワードを抽出し、かつ発生頻度及び／又は文書構造の重み付けにより優先度順に該主要キーワードのリストを含むインデックス情報を作成するインデックス情報作成手段を備えていることを特徴とする請求項1に記載の言語間情報検索支援システム。

【請求項3】 前記サーバシステムは、更に、要求された検索キーワードを用いて前記文書情報を検索する文書情報検索手段を備えており、検索された文書情報が前記要求された文書情報となるように構成されていることを特徴とする請求項1又は2に記載の言語間情報検索支援システム。

【請求項4】 前記ユーザインタフェース手段は、更に、前記検索された文書情報の検索結果リストの中で、該文書情報それぞれの対訳された主要キーワードを支援情報として表示するように構成されていることを特徴とする請求項3に記載の言語間情報検索支援システム。

【請求項5】 前記サーバシステムは、更に、前記要求された検索キーワードを前記言語間対訳辞書データベースを用いて、検索対象言語に対訳する検索キーワード対訳手段を備えており、該対訳された検索キーワードを用いて文書情報検索手段が文書情報を検索するように構成されていることを特徴とする請求項3又は4に記載の言語間情報検索支援システム。

【請求項6】 前記サーバシステムは、更に、前記検索キーワードがフレーズであれば、該フレーズから1つ以上の該検索キーワードを抽出し、かつ該検索キーワードをAND条件として前記検索キーワード対訳手段へ送る検索キーワード抽出手段を備えていることを特徴とする請求項3から5のいずれか1項に記載の言語間情報検索システム。

【請求項7】 文書情報及び該文書情報の原言語の主要キーワードを含んでいるリソース情報データベースを、更に備えていることを特徴とする請求項3から6のいずれか1項に記載の言語間情報検索支援システム。

【請求項8】 前記リソース情報データベースは、リソ

ース発見エンジンによって前記ネットワークを介して自動的に収集された前記文書情報のテキストリソースが記録されているテキストデータベースと、該文書情報それぞれに対応するインデックス情報が記録されているインデックスデータベースとを備えていることを特徴とする請求項3から7のいずれか1項に記載の言語間情報検索支援システム。

【請求項9】 前記インデックス情報作成手段が作成する前記インデックス情報は、更に、前記文書情報のID、URLパス情報、ドメイン、原言語の種類、文書タイトル、ファイル名、文書サイズ及び更新日時の少なくとも1つの項目を含んでいることを特徴とする請求項3から8のいずれか1項に記載の言語間情報検索支援システム。

【請求項10】 前記ユーザインタフェース手段は、更に、前記支援情報フレーム内に辞書引きフォームを有しており、該辞書引きフォームに入力された前記文書情報の原言語のキーワードを、前記文書主要キーワード対訳手段が所望の言語に対訳して、該対訳されたキーワードを該辞書引きフォームで表示するように該文書情報をエンハンスして応答するものであることを特徴とする請求項3から9のいずれか1項に記載の言語間情報検索支援システム。

【請求項11】 前記ユーザインタフェース手段は、更に、前記文書情報のタイトル中のキーワードを所望の言語に対訳して前記支援情報フレーム内に表示するように該文書情報をエンハンスして応答するものであることを特徴とする請求項

【請求項12】 前記ユーザインタフェース手段は、更に、前記文書情報の原言語の主要キーワードと、該原言語の主要キーワードを前記所望の言語に対訳した主要キーワードとを並べて前記支援情報フレーム内に表示するように該文書情報をエンハンスして応答するものであることを特徴とする請求項1から11のいずれか1項に記載の言語間情報検索システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、複数のコンピュータが接続されたネットワークにあって、所望の言語以外の言語で提供された文書情報の検索を支援するシステムに関するものである。

【0002】

【従来の技術】本システムの代表例としては、インターネットにおいて、膨大な数のホームページを検索するために、キーワード等の検索条件を指定して検索結果を得ることができるサーチエンジンと称される情報検索システムがある。以下では、主にインターネット上での利用を例にとり説明していく。

【0003】従来のサーチエンジンは、検索者にとって、所望の言語と検索対象のホームページの言語とが異

10

20

30

40

50

なる場合、検索者にとって適切な情報を検索をするのは非常に困難である。なぜなら、サーチエンジンで指定する検索キーワードは、少なくとも検索対象のホームページの原言語でなければならない。

【0004】このような不都合を解決できる言語間情報検索システムとして、例えば検索者の所望の言語が日本語である場合、日本語で指定された検索キーワードを自動的に英語のキーワードに対訳して、その対訳されたキーワードを用いて既存の英語のサーチエンジンで検索するシステムがある。このサーチエンジンの例としては、ドラゴンサーチ(<http://www.dragon.co.jp>)がある。

【0005】他の言語間情報検索システムの例としては、日本語に限定した上で文節程度を検索フレーズとして指定でき、その検索フレーズ内の1つ以上のキーワードを自動的に抽出してAND検索を行い、その検索結果リストの表示の際に該当するホームページのタイトルに限って日本語に翻訳して表示することができるシステムがある。このサーチエンジンの例としては、NTTのTITAN(<http://isserv.tas.ntt.jp/chisho/titan.htm>)がある。例えば検索フレーズを「日本文化として」と入力すると、検索キーワードが「Japan」及び「culture」に対訳されてAND条件で検索される。

【0006】

【発明が解決しようとする課題】前述した従来の言語間情報検索システムは、いずれも検索者が指定する検索キーワード及び検索フレーズに対して、できる限り最適な検索を行おうとするものである。従って、検索者が目的とするホームページを得るためには、確実な検索キーワード等の条件指定が必要となる。また複数のURLを表示した検索結果リストから適当なURLを選択するための判断材料は、ホームページのタイトル部等の情報しかない。更に、前述したサーチエンジンを用いて、たとえ最適なホームページを検索できたとしても、詳細なアンカーを選択することは難しい。なぜなら、検索者にとって所望の言語以外の言語で提供されたホームページは、大まかな概要を理解するのも難しい場合があるからである。

【0007】検索によって得られた当該ホームページを翻訳機に通すことで、検索者がその内容を把握することも可能である。しかしながら、現状のパーソナルコンピュータレベルで使用される翻訳機を利用しても適切な表現の翻訳効果は得られず、かつマシンの処理能力もかなり必要となる。翻訳機能をクライアント又はサーバのどちらに置いたとしても、インターネットのような膨大なマシンが接続されたオンライン上で、リアルタイムな翻訳効果を得ることは非常に困難である。

【0008】そこで本発明は、所望の言語以外の言語で提供された文書情報を検索する際に、容易かつ的確な検索を行うための支援情報を検索者に提供することを目的としている。

【0009】

【課題を解決するための手段】即ち、本発明によれば、複数のコンピュータが接続されたネットワークに設けられており、複数の種類の言語間でキーワードを対訳するための辞書を有する言語間対訳辞書データベースと、要求された文書情報の原言語の主要キーワードを言語間対訳辞書データベースを用いて所望の言語に対訳する文書主要キーワード対訳手段、及び該文書情報と同時に、対訳された主要キーワードを支援情報フレームとして表示するように該文書情報をエンハンスして応答するユーザインタフェース手段を有するサーバシステムとを備えていることを特徴とする言語間情報検索支援システムが提供される。

【0010】本発明によって、第1に、所望の言語以外の言語で提供された文書情報を閲覧すると同時に、所望の言語に対訳された主要キーワードが閲覧できるので、該文書情報の概要が把握しやすくなるという情報検索支援上の利点がある。第2に、検索対象言語の主要キーワードの情報がエンハンスされた文書情報を発信することにより、既存の情報検索システムによる複数言語のキーワードの検索を可能にするという情報発信上の利点もある。更に、文書情報の全てのテキストソースを翻訳するのに比べてキーワードしか対訳しないために、マシンの処理能力が小さくてすみ、リアルタイムな対訳が可能となる。このようなサーバを複数種類の言語間の中間物としてインターネット上に提供することにより、提供される文書情報の言語に比較的とらわれずに目的とする文書情報を探することができるようになる。

【0011】前述の「文書情報をエンハンスして」とは、元の文書情報のテキストソースに支援情報フレームのテキストソースを付加することをいう。クライアントに用意された既存のブラウザによって、2つのフレームの表示が可能となる。支援情報フレームは、例えばホームページのHTMLソーステキストのフレームタグで書くことができる。また、原言語の文書情報に他の言語の支援情報が付加されているので、既存のサーチエンジンによって該他の言語をキーワードとした検索が可能となる。

【0012】サーバシステムは、更に、文書情報のテキストソース全体から原言語の前記主要キーワードを抽出し、かつ発生頻度及び／又は文書構造の重み付けにより優先度順に該主要キーワードのリストを含むインデックス情報を作成するインデックス情報作成手段を備えてもよい。文書情報のテキストソース全体から主要キーワードが抽出されるので、予め主要キーワードを指定しておく必要もなく、内容が更新されてもより最適な主要キーワードの抽出が行える。これは、情報検索支援及び情報発信の両方において効果的である。文書構造の重み付けとは、例えばホームページのHTMLソーステキストにおいてのタグに対して重み付けしたようなものであ

る。

【0013】サーバシステムは、更に、要求された検索キーワードを用いて前記文書情報を検索する文書情報検索手段を備えてもよい。文書情報検索手段を備えることにより、既存の検索キーワードによる情報検索システムと同様に利用できるようになる。

【0014】ユーザインタフェース手段は、更に、検索された文書情報の検索結果リストの中で、該文書情報それぞれの対訳された主要キーワードを支援情報として表示するように構成してもよい。検索者は、検索結果リストが表示された段階において、それぞれの文書情報の主要キーワードが所望の言語で表示されているために、該文書情報を直接閲覧しなくても大まかな概要を把握することができるようになる。

【0015】サーバシステムは、更に、要求された検索キーワードを言語間対訳辞書データベースを用いて、検索対象の原言語に対訳する検索キーワード対訳手段を備えており、該対訳された検索キーワードを用いて文書情報検索手段が文書情報を検索するように構成してもよい。検索者が、検索対象言語の種類を指定し、検索キーワードを所望の言語で入力することによって、自動的に検索対象の原言語に対訳されて検索されるようになる。

【0016】サーバシステムは、更に、検索キーワードがフレーズであれば、該フレーズから1つ以上の該検索キーワードを抽出し、かつ該検索キーワードをAND条件として前記検索キーワード対訳手段へ送る検索キーワード抽出手段を備えてもよい。検索キーワードの指定がフレーズで行えるので、より限定した検索が可能となる。

【0017】文書情報及び該文書情報の原言語の主要キーワードを含んでいるリソース情報データベースを、更に備えていてもよい。主要キーワードが予め文書情報から抽出されかつ記録されているために、リソース情報データベースは迅速な検索結果の応答が可能となる。

【0018】リソース情報データベースは、リソース発見エンジンによってネットワークを介して自動的に収集された前記文書情報のテキストリソースが記録されているテキストデータベースと、該文書情報それぞれに対応するインデックス情報が記録されているインデックスデータベースとを備えてもよい。検索結果リストを表示する際の該文書情報の全ての情報が、インデックスデータベース内の対応するインデックス情報に記録されているために、リソース情報データベースは迅速な検索結果の応答が可能となる。

【0019】インデックス情報作成手段が作成する前記インデックス情報は、更に、前記文書情報のID、URLパス情報、ドメイン、原言語の種類、文書タイトル、ファイル名、文書サイズ及び更新日時の少なくとも1つの項目を含んでいてもよい。リソース情報データベースは、検索結果リストを表示する際にインデックスデータ

ベースのみを検索すれば、表示すべき全ての項目情報が得られるので、迅速な応答が可能となる。

【0020】ユーザインタフェース手段は、更に、支援情報フレーム内に辞書引きフォームを有しており、該辞書引きフォームに入力された前記文書情報の原言語のキーワードを、前記文書主要キーワード対訳手段が所望の言語に対訳して、該対訳されたキーワードを該辞書引きフォームで表示するように該文書情報をエンハンスして応答するものであってもよい。検索者は、閲覧している文書情報のキーワードであって支援情報フレームに表示されていない該キーワードについても、容易に対訳結果を得ることができるようになる。

【0021】ユーザインタフェース手段は、更に、文書情報のタイトル中のキーワードを所望の言語に対訳して支援情報フレーム内に表示するように該文書情報をエンハンスして応答するものであってもよい。文書情報と同時に表示されている支援情報フレームに、該文書情報のタイトルが所望の言語で表示されているので、検索者にとっては該文書情報の概要を把握するための支援情報となる。

【0022】ユーザインタフェース手段は、更に、文書情報の原言語の主要キーワードと、該原言語の主要キーワードを所望の言語に対訳した主要キーワードとを並べて支援情報フレーム内に表示するように該文書情報をエンハンスして応答するものであってもよい。検索者は、支援情報フレームに表示された原言語の主要キーワードと所望の言語の主要キーワードとを見比べることによって、所望の言語以外の言語で提供された文書情報を閲覧する際に、内容が把握しやすくなる。

【0023】

【発明の実施の形態】図1は、本発明の一実施形態におけるシステムの構成図である。以下、この構成について説明する。

【0024】クライアント13は、パーソナルコンピュータのような端末であり、一般的なWWWブラウザが用意されている。該クライアント13及びサーバはインターネット14によって接続されている。

【0025】サーバは、複数の種類の言語間でキーワードを対訳するための辞書を有する言語間対訳辞書データベース11と、文書情報を含むリソース情報データベース12と、サーバシステム10とから構成されている。該サーバシステム10は、ユーザインタフェース手段10aと、検索キーワード対訳手段10bと、文書情報検索手段10cと、文書主要キーワード対訳手段10dと、検索キーワード抽出手段10eと、インデックス情報作成手段10fとから構成されている。該リソース情報データベース12は、インデックスデータベース12aとテキストデータベース12bとから構成されている。

【0026】図2は、本発明の一実施形態におけるクラ

クライアント13とサーバシステム10との間の言語間情報検索支援のシーケンス図である。以下、各構成要素の機能と作用について、該シーケンスに沿って説明していく。

【0027】最初にリソース情報データベース12の機能と前処理について説明する。サーバシステム10は、前処理として予め検索対象となるホームページを収集しておかなければならない。これは、ロボットと称されるリソース発見エンジンによって、ネットワーク上の多くの情報リソースからホームページが自動的に収集される(S1)。収集されたホームページのテキストリソースは、リソース情報データベース12のテキストデータベース12bに記録される(S2)。但し、本サーバシステム10は、あくまで情報検索支援のために提供されるものなので、収集したホームページのHTMLテキストリソース及び該ホームページのURL及びドメイン等の各種情報のみが記録されるようにしている。

【0028】収集されたホームページがテキストデータベース12bに記録された後で、インデックス情報作成手段10fにより、該ホームページのインデックス情報が作成される(S3)。インデックス情報作成手段10fは、ホームページのテキストリソース全体から原言語の主要キーワードを抽出し、かつ発生頻度及び／又は文書構造の重み付けにより優先度順に該主要キーワードのリストを含むインデックス情報を作成する。作成された該インデックス情報は、インデックスデータベース12aへ記録される(S4)。つまり、テキストデータベース12bに記録されているテキストリソースそれぞれに対応するインデックス情報が、インデックスデータベース12aに記録されている。

【0029】このインデックス情報が、検索者に対する支援情報となる。インデックス情報には、更に、そのホームページに関する情報、即ち、ID、URLパス情報、ドメイン、言語、タイトル、サイズ及び更新日時の項目が記録されている。少なくとも、検索結果リストで表示される項目については、インデックス情報内の項目として記録されるようにする。

【0030】言語間情報検索支援のシーケンスを説明する。最初に、クライアント13において、検索者は、言語間情報検索支援サービスを受けるために、インターネット14を介してサーバシステム10へアクセスする(C1)。クライアント13の端末のユーザインタフェースには、一般のWWWブラウザが用いられている。

【0031】次に、サーバシステム10において、ユーザインタフェース手段10aが、言語間情報検索のための条件の指定を促す主ホームページをクライアント13へ転送する(S5)。

【0032】次に、クライアント13において、受け取られた主ホームページは、ブラウザに図4のように表示される(C2)。図4は、本発明の一実施形態における

検索条件指定画面である。該検索条件指定画面は、検索者に対して、検索キーワード入力部40、AND及び／又はOR条件指定部41、検索対象言語の種類の指定部42及び検索対象ドメインの指定部43への入力を促す。そこで検索者は、検索キーワード等の所望の検索条件を指定して、検索開始部44をクリックする。検索対象言語の種類は、日本語、英語、中国語及び韓国語を選択でき、検索対象ドメインは、日本、米国、中国、台湾、香港、韓国及びその他を選択できる。

【0033】次に、サーバシステム10において、検索キーワード等の検索条件を受け取ったユーザインタフェース手段10aは、これらの検索条件を検索キーワード対訳手段10bへ送る。

【0034】他の実施形態によれば、ユーザインタフェース手段10aは、検索キーワードがフレーズであってもよい。その場合、検索条件は検索キーワード抽出手段10eへ送られる。該検索キーワード抽出手段10eは、フレーズから1つ以上のキーワードを抽出し、かつ該検索キーワードをAND条件として検索キーワード対訳手段10bへ送る。

【0035】検索キーワード対訳手段10bは、言語間対訳辞書データベース11を用いて検索キーワードを検索対象言語へ対訳する(S6)。該対訳された検索キーワードと他の検索条件は、文書情報検索手段10cへ送られる。

【0036】文書情報検索手段10cは、対訳された検索キーワード等の検索条件に該当するインデックス情報を、リソース情報データベース12のインデックスデータベース12aから検索する(S7)。検索されたインデックス情報の各種項目は、検索結果リストを表示する際の支援情報として用いられる。ここでは、1つ以上のインデックス情報が検索される場合もあるし、検索されない場合もある。検索されたインデックス情報は、文書主要キーワード対訳手段10dへ送られる。

【0037】文書主要キーワード対訳手段10dは、インデックス情報内の主要キーワードのリストを、言語間対訳辞書データベース11を用いて検索者の所望の言語に対訳する(S8)。そして、該対訳された主要キーワード及びインデックス情報は、ユーザインタフェース手段10aへ送られる。

【0038】ユーザインタフェース手段10aは、検索されたインデックス情報の項目と対訳された主要キーワードとを検索結果リストのホームページとして作成して、クライアント13へ転送する(S9)。

【0039】次に、クライアント13において、受け取られた検索結果リストのホームページが、ブラウザに図5のように表示される(C3)。図5は、本発明の一実施形態における検索結果リストである。図5は、日本語を所望の言語とする検索者が、「日本語」、「英語」及び「中国語」で提供されたホームページを、「中国」、

「企業」及び「投資」という検索キーワードによって検索しようとした場合50のリストである。リスト(1)51は、文書タイトル53が「Asia Business Information Center」、主要キーワード54が「アジア、ビジネス、ニュース、開発区、進出、誘致」であることが把握できる。また、リスト(2)52は、中国語である文書タイトル55がイメージで表示されており、タイトルのみでは内容の概要さえもわからないが、主要キーワード56が「イエローページ、会社、名簿、ビジネス、サービス業」であることが把握できる。該ホームページは中国語及び英語の混在で提供されたものであるけれども、日本語に対訳された主要キーワード56が検索者にとっての支援情報となっている。クライアント13の検索者は、リスト(2)52のホームページを検索したいならば、該リストのアンカーをクリックすることで、ホームページの閲覧を要求することができる。

【0040】次に、サーバシステム10において、ホームページの閲覧要求を受け取ったユーザインタフェース手段10aは、該当するホームページのアドレスを文書情報検索手段10cへ送る。該文書情報検索手段10cは、リソース情報データベース12のテキストデータベース12bから該当するホームページのテキストリソースを検索し、これをユーザインタフェース手段10aへ戻す。該ユーザインタフェース手段10aは、検索された文書情報のテキストリソースに、対訳された主要キーワードのリスト及び辞書引きフォームを表示する支援情報フレームをエンハンスして、クライアントへ転送する(S10)。

【0041】次に、クライアント13において、受け取られたホームページが、ブラウザに図6のように表示される(C4)。図6には、本発明の一実施形態における文書情報フレーム60及び支援情報フレーム61が表されている。文書情報フレーム60が英語及び中国語が混在しているオリジナルのホームページであり、支援情報フレームが検索者に有効な支援情報である。

【0042】ここで、支援情報フレーム61の内容について説明する。まず第1に、文書タイトル62が原言語で表示される。ここでは英語で表示されている。次に、該文書タイトル中のキーワード63が所望の言語に対訳されて表示されている。ここでは日本語でキーワードの対訳が表示されている。この時、原言語のキーワードと対訳されたキーワードとを比較できるように並べて表示されており、検索者が見比べて把握できる点で有効である。次に、文書情報の主要キーワード64が所望の言語に対訳されて表示される。この時も同様に、原言語の主要キーワードと対訳された主要キーワードとを比較できるように並べて表示される。ここでは、優先順に5つの主要キーワードが表示されているが、増減することも可能である。

【0043】更に、本実施形態では、支援情報フレーム

61内に、文書情報の原言語のキーワードをオンデマンドで日本語に対訳するための辞書引きフォーム65を備えている。これは、検索者が、文書情報フレームには記載してあるが支援情報フレームには記載されていない用語を理解するための手段として提供されている。ホームページの原言語のキーワードが入力された際(C5)に、入力された原言語のキーワードを文書主要キーワード対訳処理手段を用いて所望の言語に対訳して(S11)、該対訳されたキーワードを表示する(C6)ものである。原言語の入力は、文書情報フレームのテキスト部分をマウス等でカット&ペーストで入力することができる。

【0044】図3は、本発明の一実施形態におけるクライアント13とサーバシステム10との間の言語間情報発信支援のシーケンス図である。以下、各構成要素の機能と作用について、該シーケンスに沿って説明していく。

【0045】最初に、クライアント13において、情報発信者は、ホームページのテキストリソースに支援情報をエンハンスするサービスを受けるために、インターネット14を介してサーバシステム10へアクセスする(C7)。

【0046】次に、サーバシステム10において、ユーザインタフェース手段10aが、エンハンスするホームページと検索対象言語の種類とを指定できるホームページ指定画面をクライアント13へ転送する(S12)。ここで条件指定を促すインタフェースは、特にホームページである必要はない。

【0047】次に、クライアント13において、受信されたホームページ指定に従って、エンハンスしたいホームページのテキストリソースをサーバシステム10へ送信し、検索対象言語を指定する。検索対象言語は1つである必要はなく、複数指定することができる。

【0048】次に、サーバシステム10において、エンハンスしたいテキストリソースを受け取ったユーザインタフェース手段10aは、該テキストリソースをインデックス情報作成手段10fへ送る。

【0049】インデックス情報作成手段10fは、該テキストリソース全体から原言語の主要キーワードを抽出し、かつ発生頻度及び／又は文書構造の重み付けにより優先度順に該主要キーワードのリストを含むインデックス情報を作成する(S13)。該作成されたインデックス情報は、文書主要キーワード対訳手段10dへ送られる。

【0050】文書主要キーワード対訳手段10dは、インデックス情報の主要キーワードを検索対象言語に対訳される。該インデックス情報及び該対訳された主要キーワードは、ユーザインタフェース手段10aへ送られる。

【0051】ユーザインタフェース部10aは、クライ

アントから送られたテキストリソースに、対訳された主要キーワードを含む支援情報フレーム用のテキストリソースをエンハンスする。そして、エンハンスされたテキストリソースをクライアント13へ送信する。

【0052】クライアント13において、エンハンスされたテキストリソースが受け取られる。クライアント13は、該エンハンスされたテキストリソースを元のテキストリソースと同様に情報発信することができる。該エンハンスされたテキストリソースの原言語と異なる言語を所望する検索者は、所望の言語で検索できるのはもちろん、対訳された主要キーワードの支援情報フレームを見ることが出来る。

【0053】

【発明の効果】本発明によって、第1に、所望の言語以外の言語で提供された文書情報を閲覧する際に、同時に所望の言語に対訳された主要キーワードが閲覧できるので、該文書情報の概要が把握しやすくなるという情報検索支援上の利点がある。第2に、検索対象言語の主要キーワードの情報がエンハンスされた文書情報を発信することにより、既存の情報検索システムによる検索対象言語によるキーワードの検索が可能になるという情報発信上の利点もある。更に、文書情報の全てのテキストリソースを翻訳するのに比べてキーワードしか対訳しないために、マシンの処理能力が小さくてすみ、リアルタイムな対訳が可能となり、インターネット上でのサーバに相応しいものとなる。

【0054】上述の説明では一実施形態としてインターネットを例にとり説明したが、複数のコンピュータが接続されたネットワーク上での情報検索への適用において、本発明の技術思想及び見地の範囲の種々の変更、修正及び省略は、当業者によれば容易に行うことができる。つまり、日本語、英語及び中国語等を例にとり説明したが、ロシア語等の他の言語でも同様に行えるのは当然である。従って、上述した説明は、あくまで例であって制約しようとするものではなく、本発明は、特許請求の範囲及びその等価物として限定するものだけに制約される。

【図面の簡単な説明】

【図1】本発明の一実施形態におけるシステムの構成図である。

*【図2】本発明の一実施形態におけるクライアントとサーバシステム間の言語間情報検索支援のシーケンス図である。

【図3】本発明の一実施形態におけるクライアントとサーバシステム間の言語間情報発信支援のシーケンス図である。

【図4】本発明の一実施形態におけるクライアントのブラウザに表示された検索条件指定画面である。

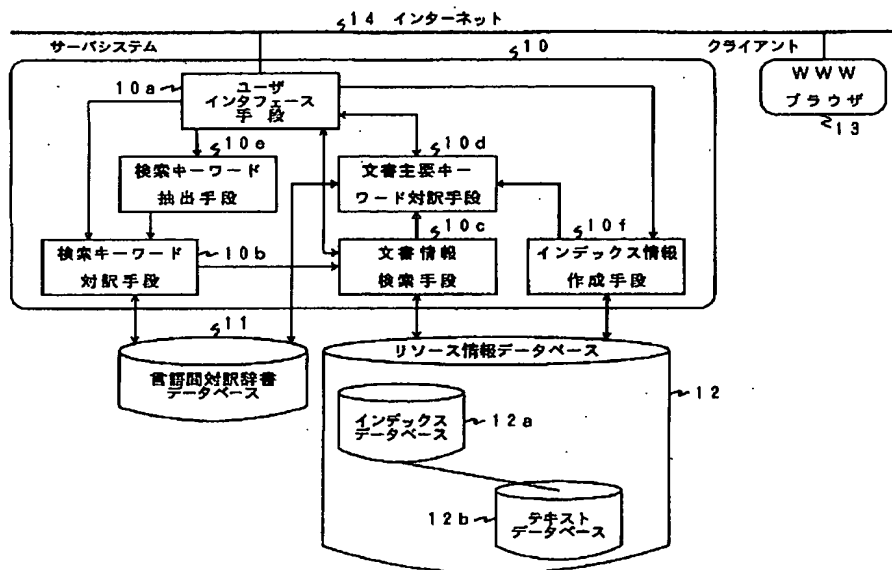
【図5】本発明の一実施形態におけるクライアントのブラウザに表示された検索結果リストである。

【図6】本発明の一実施形態におけるクライアントのブラウザに表示された文書情報フレーム及び支援情報フレームである。

【符号の説明】

- 10 サーバシステム
- 11 言語間対訳辞書データベース
- 12 リソース情報データベース
- 13 クライアント
- 14 インターネット
- 10a ユーザインタフェース手段
- 10b 検索キーワード対訳手段
- 10c 文書情報検索手段
- 10d 文書主要キーワード対訳手段
- 10e 検索キーワード抽出手段
- 12a インデックスデータベース
- 12b テキストデータベース
- 40 検索キーワードの入力部
- 41 AND及び／又はOR条件指定部
- 42 検索対象言語の種類の指定部
- 43 検索対象ドメインの指定部
- 44 検索開始部
- 50 指定された検索条件
- 51、52 リスト
- 53、55、62 文書タイトル
- 54、56、64 主要キーワード
- 60 文書情報フレーム
- 61 支援情報フレーム
- 63 文書タイトル中のキーワード
- 65 辞書引きフォーム

【図1】



【図3】

＜言語間情報発信支援シーケンス＞

[クライアント側]

[サーバシステム側]

C7)サーバシステムへのアクセス

C8)ホームページ指定画面の表示
(テキストリスト及び言語指定)

C9)エンハンスされたホームページの
テキストリストを受信

S12)ホームページ指定画面の転送。

S13)インデックス情報作成手段により要求された
テキストリストからインデックス情報を作成。

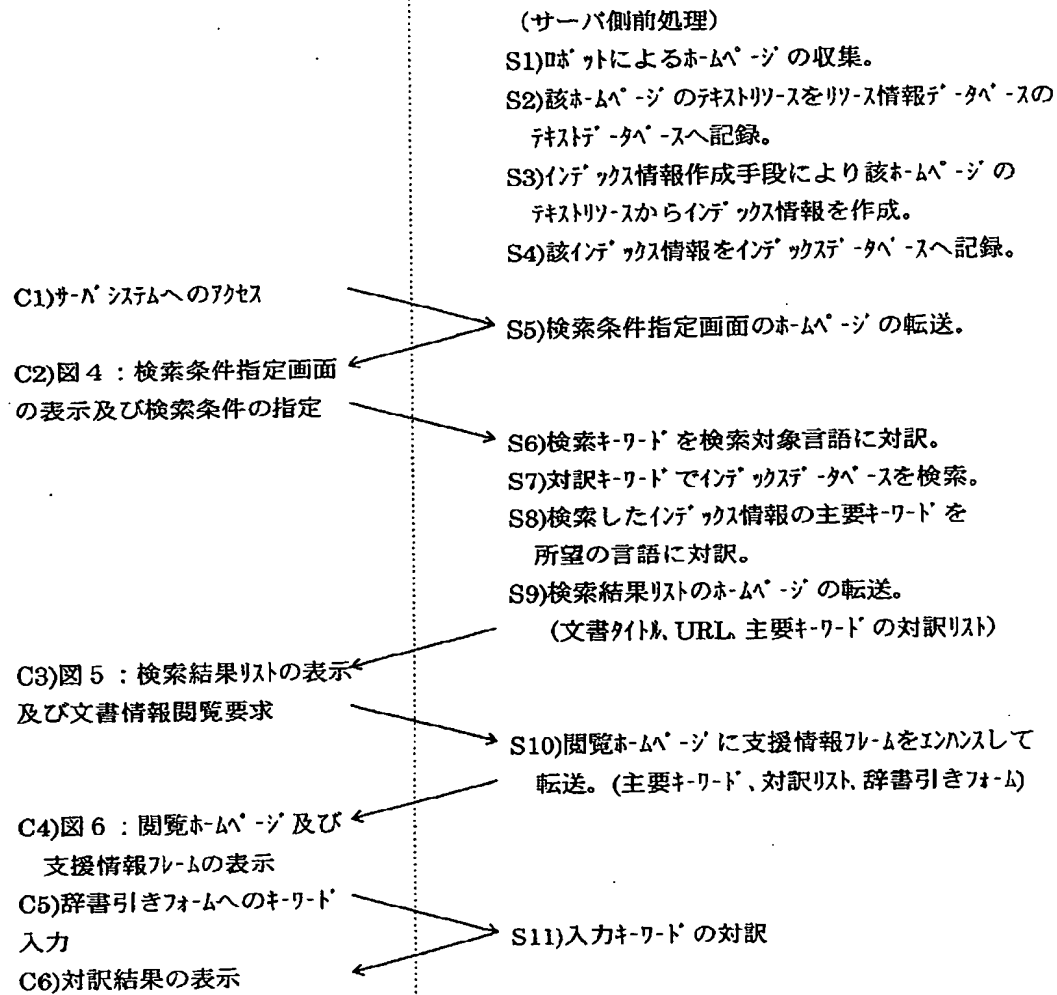
S14)ユーザインタフェース手段により該インデックス情報を
テキストリストにエンハンスして、クライアントへ送信。

【図2】

<言語間情報検索支援シーケンス>

[クライアント側]

[サーバシステム側]



【図4】

CROSS-LINGUAL INFORMATION LINKS SERVER
KDD R&D Laboratories

JAPANESE ENGLISH WHAT'S NEW GUIDE ABC

日本語または英語でキーワードを入力してください。
キーワードとキーワードの間は半角か全角のスペースを入れてください。

検索条件を指定して下さい
☐ AND ☐ OR

検索対象言語を次の中から選んで下さい
☐ 日本語 ☐ 英語 ☐ 中国語 ☐ 韓国語

検索対象ドメインを次の中から指定して下さい

☒ 日本 (jp) ☐ 米国 (comなど) ☐ 中国 (cn)
☐ 台湾 (tw) ☐ 香港 (hk) ☐ 韓国 (kr) ☐ その他

SEARCH

【図5】

検索結果一覧

検索キーワード: 中国、企業、投資
検索対象言語: 日本語/英語/中国語
ドメイン: 日本/米国/中国

Result List in Chinese(5件) Result List in English(15件) Result List in Japanese(10件)

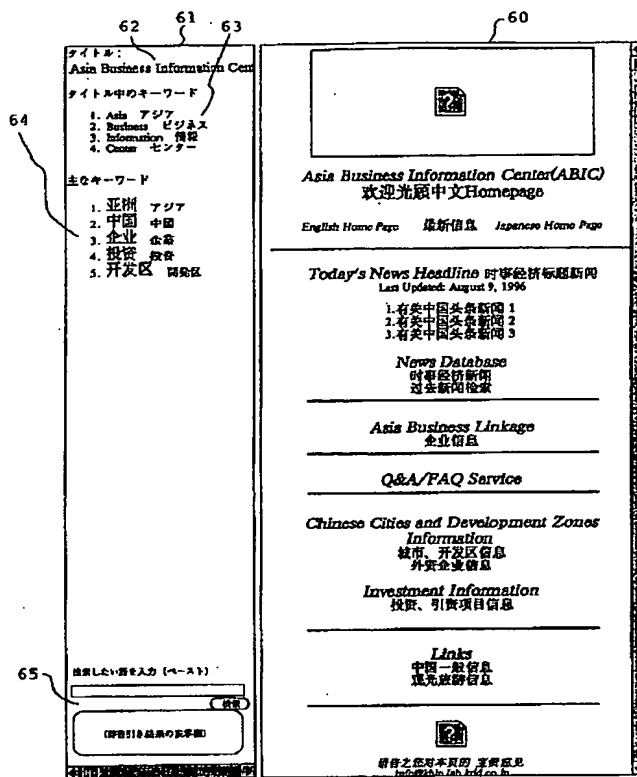
検索結果リスト

51	(1) Japan	NEW	Asia Business Information Center (Japanese SJS Display) http://www.sjispan.co.jp/kddmlb/index.html (アジア、ビジネス、ニュース、関税、通関、商社)
	45bytes (score=360)	96/10/01	
52	(2) China	RECENT	中国投資一企業名録 (Japanese SJS Display) http://china.pages.co.kr/company-guide/tydir.html (イエローページ、会社、名簿、ビジネス、サービス類、)
	510bytes (score=150)	96/09/15	

Next 10 results Previous 10 results New Search

mailto:rc@links@lab.kdd.co.jp

【図6】



フロントページの続き

(51) Int. Cl.⁶

識別記号

F I

G 0 6 F 15/401

3 1 0 A